Restricting n to Two: When Merge Requires Search

著者	後藤 亘
著者別名	Nobu GOTO
journal or publication title	Journal of business administration
volume	91
page range	105-114
year	2018-03
URL	http://id.nii.ac.jp/1060/00009633/

Restricting *n* to Two: When Merge Requires Search

Nobu Goto

Abstract
This squib aims to increase the simplicity and efficiency of the system of language by optimizing Merge by search and minimizing search by labels at a phase level. More specifically, in terms of computational efficiency and minimization of working memory, this squib proposes that Merge requires search to optimize its application in conformity to *n*=2 and search requires labels to minimize its application in working memory. The proposed system integrates Chomsky's (2007, 2008) phase theory and Chomsky's (2013, 2015) labeling theory in that the input of Merge (*n*) is optimized at the phase level (CP and v*P) in collaboration with labeling and search. As one consequences of the proposed system, it will be illustrated that parametric differences of Condition on Extraction Domains (CED) in English and Japanese that have resisted a satisfactory explanation in the minimalist literature on extraction receive a straightforward explanation under the view that an unlabeled {XP, YP} is opaque for extraction.

1 *n*'s Problem
Merge, a primitive operation to ensure the discrete infinity of human language, takes two syntactic objects (SOs) as its input (*n*=2) and forms the set of the two SOs as its output (Chomsky 2007, 2008, 2013, 2015 among others):

(1) Merge(X, Y) = {X, Y}

Given the definition of Merge (1), two puzzles arise after and before its application:

(2) *Puzzle 1*: How to determine the label of the output of Merge {X, Y}?
Puzzle 2: How to determine the input (*n*) of Merge (X, Y)?

Discussions have been widely held on the puzzle 1 ("Problems of Projection") (Chomsky 2013, 2015 among others) but little is known about the puzzle 2, which I will call "*n*'s problem." Only with the definition of (1), it is not clear why n is X and Y, but not others (say, Z, W, *etc.*), and we cannot answer the question of so-called the problem of selection. The simplest definition of Merge, thus, cannot say anything about the process of (3):

(3) $n>2 \rightarrow n=2$

Taking into account of the fact that the process of (3) is crucially involved, for example, in extraction of SOs from the lexicon by External Merge (EM) (see Chomsky 2013:41) and extraction of an SO from a complex SO by Internal Merge (IM) in the course of the derivation (see Chomsky 2008:147 and footnote 2), it is important to clarify the process of n to state the definition of Merge more explicitly.

2 Language as an Organic System

To solve the n's problem, I develop a system in which Merge, search, and labels can interact with each other so as to be able to enhance the computational efficiency.

2.1 Optimization of Merge by Search

In terms of the computational efficiency, I first assume (4) (Goto 2016):

(4) Merge requires search to optimize its application in conformity to $n=2$.
   a. <u>When Merge does not require search: $n=2$</u>
      Merge(SO_1, SO_2) = {SO_1, SO_2}
   b. <u>When Merge requires search: $n>2$</u>
      Merge(SO_1, …, SO_n) →
      Search(SO_1, …, SO_n) → (SO_1, SO_2) →
      Merge(SO_1, SO_2) = {SO_1, SO_2}

That is, when $n=2$ Merge does not require search and applies freely without recourse to search, but it requires search when $n>2$ to optimize its application in conformity to $n=2$. To put it differently, given $n>2$, Merge tries to reduce its computational cost/load by restricting its input to two with the help of search. In the proposed system, therefore, search plays an important role in providing the simplest input for Merge. The application of Merge is always optimized by search iff $n>2$.

2.2 Minimization of Search by Labels

In terms of the minimization of working memory, I then adopt (6) (Goto 2016), assuming (5) for the role of labels (see Chomsky 2000:132 and Chomsky

2005:14 among others):

(5) In narrow syntax (NS), labeled SOs are visible but unlabeled SOs are invisible.

(6) Search requires labels to minimize its application in working memory and only labeled SOs are visible/accessible to search.

That is, search does not apply blindly. It itself operates minimally under the visibility of SOs defined by labeling (cf. footnote 5). For a different approach to (6), see Goto (2016).

Note that the proposed system is compatible with the theory of unconstrained Merge, since it only says that search, but not Merge, is sensitive to labeling: Merge can freely apply to SOs, whether labeled or unlabeled, but search is affected by labeling of SOs in NS. In the proposed system, therefore, search is directly related to labeling and labeling works to minimize the application of search in working memory.(1) The whole picture of the proposed system is thus as follows:

(7) Labels minimize search and search optimizes Merge, all interacting organically for computational simplicity and efficiency.

If this design of language is on the right track, it follows that labeling of SOs is a prerequisite for search but not for Merge, which sheds new light on the issue of the necessity of labeling in NS. While Chomsky (2007, 2008) argues that labeling of SOs is a prerequisite for applications of NS operations, Chomsky (2013, 2015) argues that labeling of SOs is not a prerequisite for applications of NS operations. Among others, the underlying rationale of the former is that labeling is necessary for the derivation to proceed and for c(ategorial)-selection to take place. Pesetsky (1982) claims that c-selection are reducible to s(emantic)-selection, but the issue seems unsettled. For arguments that c-selection cannot be subsumed under s-selection, see Odijk (1997) among others. On the other hand, the underlying rationale of the latter is that “if labeling is a prerequisite for entering into computation, it would block many cases of EM, e.g., Merge (Z, {XP, YP}) = {Z, {XP, YP}}” (Chomsky 2013:43, footnote 30). Importantly, the proposed system developed in this squib can equally acknowledge both sides of the argument, as it does not totally exclude the necessity of labeling in NS in favor of Chomsky’s (2007, 2008) approach and still keeps to the theory of unconstrained Merge in favor of Chomsky’s (2013, 2015) approach.

2.3 Correlation between Labeling and Extraction

As one consequences of the proposed system, (8) is deduced (see Goto 2016 for a different characterization of (8)):

(8) Extraction from an unlabeled SO is disallowed.

This is because extraction out of a complex SO by IM in the course of the derivation falls into $n>2$; and when $n>2$ Merge requires search to optimize its application; but unlabeled SOs are invisible/inaccessible to search; and therefore extraction from an unlabeled SO is disallowed.(2)

With respect to labeling, I adopt Chomsky's (2013, 2015) labeling theory (9) in English and Saito's (2013, 2014) labeling theory (10) in Japanese.

(9) Chomsky's (2013, 2015) labeling theory (English)
 a. SO = {H, XP}, H a head and XP not a head, is labeled (H).
 b. SO = {XP, YP}, neither a head, is not labeled and remains unlabeled.
 c. SO = $\{XP_{[F]}, YP_{[F]}\}$, F a prominent feature, is labeled (F) by feature sharing.

(10) Saito's (2013, 2014) labeling theory (Japanese)
 a. SO = {H, XP}, H a head and XP not a head, is labeled (H). (= (9a))
 b. SO = {XP, YP}, neither a head, is not labeled and remains unlabeled. (= (9b))
 c. SO = $\{XP_{\text{-K}}, YP\}$, K a structural Case, is labeled (Y), rendering XP invisible.(3)

3 Consequences

Below, I demonstrate that CED effects in English and Japanese receive a principle explanation, and suggest that the proposed analysis has far-reaching consequences for other extraction phenomena.

3.1 Explaining CED Effects

Under the present assumptions, parametric differences between English and Japanese with respect to CED effects receive a principled explanation under the view that the unlabeled {XP, YP} is opaque for extraction:

(11) CED effects in English and Japanese:
 In English, while extraction out of complements is allowed, extraction out of subjects and adjuncts (non-complements) is disallowed (Ross 1967, Huang 1982, Chomsky 1986 among others) (see (12)). On the

other hand, in Japanese, while extraction out of subjects and complements (arguments) is allowed, extraction out of adjuncts is not allowed (Saito 1985, Takahashi 1994, Ishii 1997 among others) (see (13)).

(12) *English*

a. Who$_i$ did you believe [that John saw t_i]?

b. *?Who$_i$ did John get jealous [before I talked to t_i]?

c. *?Who$_i$ did [pictures of t_i] please you?

(13) *Japanese* (NB: the awkwardness of (13a) and (13c) is due to the complex NP constraint; see Saito 1985)

a. ??Dono hon-o$_i$ Mary-ga [John-ga t_i kat-ta koto]-o
which book-acc M.-nom J.-nom buy-past fact-acc
mondai-ni siteru no
no problem-dat making Q
'Which book is it that Mary is calling the fact that John bought into question?'

b. *Sono hon-o$_i$ John-ga [minna-ga t_i kau node]
that book-acc J.-nom all-nom buy because
tigau hon-o kat-ta
different book-acc buy-past
'Because everyone buys that book, John bought a different one.'

c. ??Dono hon-o$_i$ Mary-ga [John-ga t_i kat-ta koto]-ga
which book-acc M.-nom J.-nom buy-past fact-nom
monadi-da to omotteru no
problem-is comp think Q
'Which book does is it that Mary thinks the fact that John bought it is a problem?'

In the proposed system, CED effects in English (12) are explained as follows: extraction out of complements {H, XP} is allowed because they are labeled and visible to search, whereas extraction out of non-complements {XP, YP} is disallowed because they are not labeled and invisible to search. On the other hand, CED effects in Japanese (13) are explained as follows: extraction out of arguments is allowed because they are Case-marked and visible to search, whereas extraction out of adjuncts is disallowed because they are not Case-marked and invisible to search. Significantly, the proposed analysis of subject-extraction is perfectly compatible with Chomsky's (2008) in-situ analysis in that it must be from the base position, not from the surface position: the base structure {DP, v*P} is of the form {XP, YP} that is not labeled, as noted in Chomsky (2013, 2015), and hence extraction out of the unlabeled SO is disallowed (for further supporting evidence for the in-situ analysis of subject-extraction, see Müller 2011:104).(4)

Regarding wh-extraction from complements, one might wonder why unlabeled {XP, YP} structures occupying SPEC-v* such as {DP, v*P} (subject) and {v*P, CP} (adjunct) do not block wh-extraction out of complements. However, it is important to notice that in the phase-based derivational system (Chomsky 2007, 2008), v*-search can optimize the derivation for further computations at the CP-phase level in terms of the Phase-Impenetrability Condition (PIC). That is, since the complement {V, XP} resides in the interior of the phase v* head, v*-search can optimize the derivation for the CP-phase level by searching into the labeled/visible SO. Note that in the proposed system, since search works to enable Merge to apply simply in conformity to n=2, it follows that the derivation at the CP-phase level has already been optimized at the v*P-phase level, so that Merge at the CP-phase level does not call for search, being able to satisfy n=2, with no help of search.(5)

In relation to this, it is also worth noting that v*-search cannot optimize the derivation with the adjunct island effect. This is because the adjunct XP is outside of v*-search, as in {{v* …}, XP}. Hence the adjunct {XP, YP} structure is not labeled and remains invisible to search at the CP-phase level. As a result, extraction from the unlabeled SO is disallowed. A potential question of this analysis of the adjunct island effect is how adjuncts get their labels in time to be sent to the interfaces. However, given that unlabeled adjuncts are interpreted as a modifier at the semantic interface, as proposed by Hornstein and Pietroski (2009), it follows that adjuncts need not to be necessarily labeled in NS, but do not crash at the interfaces.

3.2 Prospects of the Proposed System

The proposed system predicts:

(14) An unlabeled {XP, YP} structure is (basically) opaque for extraction but becomes transparent if it is labeled either by (9c) or (10c) (or some other means).

Above I have argued that this prediction is borne out by the CED effects in English and Japanese.

Note that the prediction is not limited to them. Arguably, *Complex XP Constraint* (Bošković 2015) (including *Complex NP constraint* in the sense of Ross 1967), *Coordinate Structure Constraint* (Ross 1967), *Freezing Effects*

(Wexler & Culicover 1980), *Proper Binding Condition* (PBC) (Fiengo 1977), *etc.* can all be unified under the same view. On closer inspection, it turns out that domains that do not allow extraction are of the form {XP, YP} that is not labelable, whereas domains that allow extraction are of the form {XP, YP} that is labelable:

(15) a. $^{*}XP_i \ldots \boxed{[_{\alpha}} \ldots t_i \ldots]$ (α is not labeled, which is shaded with gray)
b. $^{\checkmark}XP_i \ldots \boxed{[_{\beta}} \ldots t_i \ldots]$ (β is labeled)

In the proposed system, extractability is predictable by labelability at the phase level. In a manner of speaking, we may be able to take the current attempt as reviving Chomsky's (1986) framework, as unlabeled SOs work equivalent to "barriers" in NS (see Goto 2016 for further empirical support of this analysis).

4 Conclusion

In this squib, with the aim of increasing the simplicity and efficiency of the system of language, I have proposed that Merge requires search to optimize its application in conformity to *n*=2 and search requires labels to minimize its application in working memory. The proposed system can be regarded as an integrated system of Chomsky's (2007, 2008) phase theory and Chomsky's (2013, 2015) labeling theory, in that search, which is sensitive to labeling, plays a crucial role in determining the input of Merge (*n*) at the phase level. As one consequences of the proposed system, I have illustrated that the CED effects in English and Japanese receive a straightforward explanation under the view that the unlabeled {XP, YP} is opaque for extraction but becomes transparent if it is labeled either by (9c) or (10c), and suggested that this view has far-reaching consequences for other extraction phenomena that had so far been treated by separate conditions.

In any case, further investigation is necessary, and even if the overall picture of the system developed in this squib turns out to be wrong, the "*n*'s problem" still remains and demands close investigation insofar as *n*=2 is required for the definition of the simplest Merge.

References

Bošković, Željko. 2015. Deducing the generalized XP constraint from phasal spell-out, Ms., University of Connecticut.

Chomsky, Noam. 1986. *Barriers*. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. by Roger

Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36, 1-22.

Chomsky, Noam. 2007. Approaching UG from below. In *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax-Semantics*, ed. by Uli Sauerland and Hans-Martin Gärtner, 1-29, Berlin.: Mouton de Gruyter.

Chomsky, Noam. 2008. On phases. In *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, ed. by Robert Freiden, Carlos Peregnn Otero & Maria Luisa Zubizarreta, 133-166. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130, 33-49.

Chomsky, Noam. 2015. Problems of projection: Extensions. In *Structures, strategies and beyond-studies in honor of Adriana Belletti,* ed. by Elisa Di Domenico, Cornelia Hamann & Simona Matteini, 3-16. Amsterdam/Philadelphia: John Benjamins.

Epstein, Samuel D., Hisatsugu Kitahara, and T. Daniel Seely. 2014b. *What do we wonder is not syntactic?* Unpublished manuscript, University of Michigan Keio University Eastern Michigan University.

Fiengo, Robert. 1977. On trace theory. *Linguistic Inquiry* 8, 35–61.

Goto, Nobu. 2016. Labelability = extraetability: It's theoretical implications fot the free-Merge hypothesis. *Proceedings of NELS 46,* 335-348.

Hornstein, Nobert and Pietroski, Paul. 2009. Basic operations: Minimal syntax-semantics. *Catalan Journal of Linguistics* 8, 113-139.

Huang, C.-T. James. 1982. *Logical relations in Chinese and the theory of grammar*. Doctoral dissertation, MIT.

Ishii, Toru. 1997. *An asymmetry in the composition of phrase structure and its consequences*. Doctoral dissertation, University of California, Irvine.

Müller, Gereon. 2011. *Constraints on displacement: A phase-based approach*. John Benjamins Pub Co.

Odijk, Jan. 1997. C-selection and s-selection. *Linguistic Inquiry* 28, 365-371.

Pesetsky, David. 1982. *Paths and categories*. Doctoral dissertation, MIT, Cambridge, Mass.

Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, Mass.: MIT Press.

Ross, John R. 1967. *Constraints on variables in Syntax*. Doctoral dissertation, MIT, Cambridge, Mass.

Saito, Mamoru. 1985. *Some asymmetries in Japanese and their theoretical implications*. Doctoral dissertation, MIT, Cambridge, Mass.

Saito, Mamoru. 2013. Case and labeling in a language without φ-feature agreement. Ms., Nanzan University.

Saito, Mamoru. 2014. Kukoozoo-ni okeru bunpookaku-no yakuwari [The role of case in phrase structure]. Paper presented at Keio Linguistic Colloquium. September 1-5.

Stepanov, Arthur. 2007. The end of CED? Minimalism and extraction domain. *Syntax* 10, 80-126.

Takahashi, Daiko. 1994. *Minimality of movement.* Doctoral dissertation, University of Connecticut, Storrs.

Uriagereka, Juan. 1999. Multiple spell-out. In *Working minimalism*, ed. by Samuel Epstein and Norbert Hornstein, 251-282. Cambridge, Mass.: MIT Press.

Uriagereka, Juan. 2012. *Spell-out and the minimalist program*. Oxford: Oxford University Press.

Wexler, Ken and Peter Culicover. 1980. *Formal principles of language acquisition.* Cambridge, Mass.: MIT Press.

---

(1) Chomsky has used the term "search" in the process of Agree (probe-goal search) (Chomsky 2000) and feature sharing in the labeling algorithm (Chomsky 2013, 2014). However, it is not clear, at least to me, whether a process of searching is indeed involved in the relevant procedures. It seems to me that Agree is just a spontaneous process in the workspace under the concept of "match," and it is an automatic process that can take place under asymmetrical information about feature composition between two lexical items (LIs), which can be derived from configurations and properties of LIs (see Chomsky 2000:122). What is significant is that this kind of automatic process is not specific to language: see Alan Turing's reaction-diffusion model for pattern formation or Stuart Kauffman's self-organization theory. All else being equal and if Agree is just such a process, as part of the cognitive system, it would not be implausible to claim that search is not involved in Agree. In relation to this, if structurally asymmetrical information is sufficient for determining the label of {H, XP}, the notion of minimal search in labeling would also be irrelevant. Thus, it would be plausible to conclude that search is not involved in labeling, either. Given this, a potential problem of how to establish a probe-goal relation between T and DP in SPEC-v* by penetrating the unlabeled {DP, v*P} does not arise. This is simply because search is not involved in Agree. In this squib, I will use the term "search" in relation to the application of IM, concentrating on extraction from a complex SO by IM and abstracting away from extraction from the lexicon by EM.

(2) The observation that extraction by IM in the course of the derivation falls into $n>2$ will be easier to imagine by considering, for example, that {pictures {of {who}}} provides three candidates to the application of Merge(C): Merge(C, {pictures of who}), Merge(C, {of who}), or Merge(C, {who}). Probably, when search tries to pick up one SO for one functional head (SO), the content of features will be relevant to the process of $n>2 \rightarrow n=2$. Needless to say, further research is necessary.

(3) Epstein, Kitahara, and Seely (EKS) rationalize (10c) by reasoning that in Japanese, each overt Case particle constitutes an independent head, and after valuation, it becomes a purely phonological head which has nothing to do with NS and the Conceptual-Intentional (CI) interface. Then they claim that such purely phonological

heads cannot serve as a label-identifier at CI, and thus the label of {XP-K, YP} results in Y. Here, I adopt (9) and (10), though they should be unified, if possible.

(4) Just saying that a complex SO in the specifier constitutes an island for extraction (Uriagereka 1999) or a moved SO constitutes an island for extraction (Stepanov 2007) is insufficient to unify CED effects in (11): the former is too strong to derive them in languages like Japanese and the latter is too strong to derive them in languages like German, in which extraction out of non-moved subjects is in fact not allowed (see Müller 2011 and Uriagereka 2012 for further arguments against Stepanov 2007).

(5) It may be more relevant to state that search halts at an unlabeled SO, and from there it tries to pick up one SO from the complex SO in conformity to n=2, rather than just claiming that unlabeled SOs are invisible to search (see (6)). Given this, it follows that the inside of an unlabeled SO is still invisible but its sister is visible to search. This view will be able to accommodate wh-extraction from complements as follows: on its way to SPEC-C, a wh-phrase stops by SPEC-v*, yielding an unlabeled {Wh, v*P} structure at the edge, but the sister of the SO is visible to search, so that further extraction of the wh-phrase is freely allowed. In any case, the ideas that search is sensitive to labeling and labeling defines minimality are maintained. Thus, cases like wh-island effects can also be deduced from a minimality condition, as in Rizzi (1990).

（2018 年 1 月 5 日受理）